On weak lumpability of denumerable Markov chains

James Ledoux

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET AUTOMATIQUE

On Weak Lumpability of Denumerable Markov Chains

James Ledoux

N° 2221

Février 1994

PROGRAMME 1

Architectures parallèles,

bases de données,

réseaux et systèmes distribués

Rapport de recherche

ISSN 0249-6399

On Weak Lumpability of Denumerable Markov Chains

James Ledoux *

Programme 1 — Architectures parallèles, bases de données, réseaux et systèmes distribués
Projet Model

Rapport de recherche n˚2221 — Février 1994 — 22 pages

Abstract: We consider weak lumpability of denumerable discrete or continuous time Markov chains. Firstly, we are concerned with irreducible recurrent positive and R-positive Markov chains evolving in discrete time. We study the properties of the set of all initial distributions of the starting chain leading to an aggregated homogeneous Markov chain with respect to a partition of the state space. In particular, the asymptotic interpretation of the quasi-stationary distribution is addressed and it is fruitfully used for weak lumpability of R-positive Markov chains. Furthermore, we present a simple example which shows that a denumerable Markov chain can be (weakly) lumped into a finite Markov chain. Finally, it is stated that weak lumpability for any continuous time Markov chain with an uniform transition semi-group can be handled in discrete time context. The sequel of this result are also discussed for irreducible positive-recurrent or λ-positive continuous time Markov chains.

Key-words: Weak Lumpability, Positive recurrence, R-positivity, Quasi-stationary distribution, Uniform transition semi-group.

(Résumé : tsvp)

This work was partly supported by the grant 290C2010031305061 of the French Region Bretagne

* {ledoux}@irisa.fr

Unité de recherche INRIA Rennes
IRISA, Campus universitaire de Beaulieu, 35042 RENNES Cedex (France)
Téléphone : (33) 99 84 71 00 – Télécopie : (33) 99 38 38 32

Sur l'agrégation faible des chaînes de Markov discrètes

Résumé : Nous nous intéressons à la propriété d'agrégation faible de chaînes de Markov à espace d'état dénombrable en temps discret ou continu. Tout d'abord, nous considérons les chaînes irréductibles récurrentes positives et R-positives évoluant en temps discret. Nous étudions les propriétés de l'ensemble de toutes les distributions initiales permettant de conserver la propriété markovienne lors d'une agrégation de l'espace d'état d'une chaîne de Markov. En particulier, l'interprétation asymptotique d'une distribution quasi-stationnaire est mise en évidence et est utilisée pour l'agrégation faible de chaînes de Markov R-positives. Nous présentons également une exemple très simple qui montre qu'une chaîne de Markov à espace d'état dénombrable peut être (faiblement) agrégée en une chaîne markovienne à espace d'état fini. Finalement, il est montré que la proriété d'agrégation faible pour toute chaîne de Markov à temps continu possédant un semi-groupe de transition uniforme peut être traitée dans un contexte temps discret. Les conséquences de ce résultat général sont également signalées pour les chaînes de Markov à temps continu récurrentes positives et λ-positives.

Mots-clé : Agrégation faible, Récurrence positive, R-positivité, Distribution quasi-stationnaire, Semi-groupe de transition uniforme

1 Introduction

Let us consider a homogeneous Markov chain X, in discrete or continuous time, on a countably infinite state space denoted by E, which without loss of generality we assume to be a subset of the natural numbers $\mathbb{N}$ (i.e. $E \subseteq \mathbb{N}$.) Let $\mathcal{B} = \{B(0), B(1), \ldots\}$ be a fixed partition of E. We associate with the given chain X the aggregated chain Y, over the state space $F = \{0, 1, \ldots\}$, defined by:

$$Y_t = l \iff X_t \in B(l), \quad \text{for any } t.$$

We are interested in the set of all initial distributions of X which give an aggregated homogeneous Markov chain Y. If this set is not empty, we say that the family of Markov chains sharing the same transition semi-group is *weakly lumpable*. Most of the literature on lumpability has been concerned with the *strong lumpability* situation, that is, when any initial distribution leads to an aggregated homogeneous Markov chain. To the best of my knowledge, the weak lumpability problem with countably infinite state space has been addressed only recently in Ball and Yeo [1] for (irreducible positive-recurrent) continuous time Markov chains. The purpose of this note is to propose some results in discrete or continuous time, prolonging the studies reported in Rubino and Sericola [14],[15],[16] and Ledoux and al. [9] for a finite state space. Section 2 deals with discrete time Markov chains and mainly concerns weak lumpability for irreducible positive-recurrent or R-positive chains. In particular, we discuss the ergodic interpretation of the quasi-stationary distribution. The third section shows that lumpability for any denumerable continuous time Markov chains with an uniform transition semi-group can always be replaced in the discrete time context. The sequel of this result are also discussed for irreducible positive-recurrent or λ-positive continuous time Markov chains.

By convention, vectors are row vectors. Column vectors are indicated by means of the transpose operator $(.)^*$. The vector with all its components equal to 1 (resp. 0) is denoted merely by 1 (resp. 0). The set of all probability distributions on E will be denoted by $\mathcal{A}$. For any subset B of E and $\alpha \in \mathcal{A}$, the restriction of α to B, i.e. the vector $(\alpha(i), i \in B)$, is denoted by α_B; if $\alpha_B 1^* \neq 0$, α^B is the vector defined by $\alpha^B(i) = \alpha(i)/\sum_{j \in B} \alpha(j)$ if $i \in B$ and by 0 if $i \notin B$.

2 Weak lumpability in discrete time

Let $X = (X_n)_{n \geq 0}$ be a homogeneous Markov chain over state space E, given by its transition probability matrix $P = (P(i,j))_{i,j \in E}$ and its initial distribution α; when necessary we denote it by (α, P). Let $P(i, B)$ denote the transition probability of moving in one step from state i to the subset B of E, that is $P(i, B) = \sum_{j \in B} P(i,j)$. We denote the aggregated chain constructed from (α, P) with respect to the partition $\mathcal{B}$ by $agg(\alpha, P, \mathcal{B})$.

Definition 2.1 *A sequence $(B_0, B_1, \ldots, B_j)$ of classes of $\mathcal{B}$ is called possible for the initial distribution α iff $\mathbb{P}_\alpha(X_0 \in B_0, X_1 \in B_1, \ldots, X_j \in B_j) > 0$. Given any distribution $\alpha \in \mathcal{A}$ and a possible sequence $(B_0, B_1, \ldots, B_j)$ for α, we can define the vector $f(\alpha, B_0, B_1, \ldots, B_j) \in \mathcal{A}$ recursively by:*

$$\begin{aligned} f(\alpha, B_0) &= \alpha^{B_0} \\ f(\alpha, B_0, B_1, \ldots, B_k) &= (f(\alpha, B_0, B_1, \ldots, B_{k-1})P)^{B_k}. \end{aligned}$$

For any $B \in \mathcal{B}$, $\mathcal{A}(\alpha, B)$ denotes the subset of all distributions of the form $f(\alpha, B_1, \ldots, B_j, B)$.

By definition, the aggregated chain $Y = agg(\alpha, P, \mathcal{B})$ is a homogeneous Markov chain if and only if $\forall l, m \in F$, $\forall n \geq 0$ and $\forall (B_0, B_1, \ldots, B_{n-1}, B(l))$ possible for α,

$$\begin{aligned} & \mathbb{P}_\alpha(X_{n+1} \in B(m) \mid X_n \in B(l), X_{n-1} \in B_{n-1}, \ldots, X_0 \in B_0) \\ = \; & \mathbb{P}_\alpha(X_{n+1} \in B(m) \mid X_n \in B(l)) \end{aligned} \tag{1}$$

and the probability in the right-hand side does not depend on n; in that case, it describes the probability of going from state l to state m in one step for the aggregated chain $agg(\alpha, P, \mathcal{B})$. The approach developed in Kemeny and Snell (1976) and in [14] consists in rewriting the above conditional expression as

$$\mathbb{P}_\beta(X_1 \in B(m)) \qquad \text{with } \beta = f(\alpha, B_0, \ldots, B(l)),$$

that is, in including the past into the initial distribution. Therefore, it becomes straightforward to state the following necessary and sufficient condition for Y to be a homogeneous Markov chain without any particular assumption on X.

Theorem 2.2 *The chain $Y = agg(\alpha, P, \mathcal{B})$ is a homogeneous Markov chain iff $\forall l, m \in F$, the probability $\mathbb{P}_\beta(X_1 \in B(m))$ is the same for every $\beta \in \mathcal{A}(\alpha, B(l))$. This common value is the transition probability for the chain Y to move from state l to state m.*

The aim of this section is to study the properties of the set of distributions

$$\mathcal{A}_\mathcal{M} = \{\alpha \in \mathcal{A} / agg(\alpha, P, \mathcal{B}) \text{ is a homogeneous Markov chain } \}.$$

2.1 Weak lumpability for positive-recurrent Markov chains

Throughout this subsection, we assume that the considered Markov chain is irreducible positive-recurrent. Therefore, there exists an unique probability vector, denoted by π, which satisfies $\pi P = \pi$. Let g be a real function on E and m a probability measure on E; g is m-integrable if

$$m(|g|) \triangleq \sum_{i \in E} m(i)|g(i)| = \mathbb{E}_m[\,|g|\,] < \infty.$$

For such a Markov chain, we have the following standard corollary of the ergodic theorem.

Result 2.3 *For any bounded real function g on E, we have for all $\alpha \in \mathcal{A}$*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}_\alpha[g(X_k)] = \pi(g).$$

We only need the following lemma to derive Theorem 2.5 from Theorem 2.2 with similar arguments as for [14, Th. 3.5].

Lemma 2.4 *Let β_n be the vector $(1/n)\sum_{k=1}^{n} \alpha P^k$. For any $l, m \in F$, we have*

$$\begin{aligned} \lim_{n\to\infty} f(\beta_n, B(l)) &= f(\pi, B(l)), \\ \lim_{n\to\infty} \mathbb{P}_{f(\beta_n, B(l))}(X_1 \in B(m)) &= \mathbb{P}_{f(\pi, B(l))}(X_1 \in B(m)). \end{aligned}$$

Proof. To obtain the first limit, it suffices to let respectively $\forall i \in E : g_j(i) = 1_{\{j\}}(i)$ with any $j \in B(l)$ and $g(i) = 1_{B(l)}(i)$ in Result 2.3. Since $f(\beta_n, B(l))(j) = \sum_{k=1}^{n} \mathbb{P}_\alpha(X_k = j)/\sum_{k=1}^{n} \mathbb{P}_\alpha(X_k \in B(l))$, the numerator and the denominator tend respectively to $\pi(g_j) = \pi(j)$ and to $\pi(g) = \pi_{B(l)} 1^*$.

The second limit is derived from the previous considerations and from Result 2.3 letting $g(i) = 1_{B(l)}(i) P(i, B(m))$, for $i \in E$. Indeed, we can write

$$\mathbb{P}_{f(\beta_n, B(l))}(X_1 \in B(m)) = \frac{1}{\sum_{i \in B(l)} \beta_n(i)} \sum_{i \in B(l)} \beta_n(i) P(i, B(m)), \tag{2}$$

and the two factors in the right-hand side of formula (2) tend respectively to $1/\pi_{B(l)} 1^*$ and $\sum_{i \in B(l)} \pi(i) P(i, B(m))$. □

Finally, we have

Theorem 2.5 *If $\mathcal{A}_\mathcal{M} \neq \emptyset$, then $\pi \in \mathcal{A}_\mathcal{M}$ and the transition probability matrix of the homogeneous Markov chain $agg(\alpha, P, \mathcal{B})$, denoted by $\widehat{P}$, is the same for all $\alpha \in \mathcal{A}_\mathcal{M}$. The entries of matrix $\widehat{P}$ are given by*

$$\widehat{P}(l, m) = \sum_{i \in B(l)} \pi^{B(l)}(i)\ \ P(i, B(m)) \quad l, m \in F.$$

With the previous result, Theorem 3.7 from [14] can be extended to our denumerable context. Consequently, the set $\mathcal{A}_\mathcal{M}$ is the (a priori) infinite intersection of a decreasing sequence of convex sets, denoted by $\mathcal{A}^j$ ($j \geq 1$), which are the solutions to the linear systems defined as follows.

- For each $l \in F$, let us compute the matrix

  $$\widetilde{P}_l = \left(P(i, B(k))\right)_{i \in B(l), k \in F}$$

  and let us denote by $\widehat{P}_l$ the lth row of the transition probability matrix $\widehat{P}$. We can define, for each $l \in F$, the matrix

  $$H_l = \widetilde{P}_l - 1^* \widehat{P}_l.$$

- For each $l \in F$, let us denote by P_l the submatrix P constituted by the transition probabilities from the states of $B(l)$ to the states of E, i.e.

$$P_l = \left(P_{B(l)B(0)} \cdots P_{B(l)B(k)} \cdots \right).$$

- Let us define the block diagonal matrices

$$\begin{aligned} H^{[1]} &= \text{diag}(H_l), \\ H^{[j+1]} &= \text{diag}(P_l\, H^{[j]}), \quad j \geq 1. \end{aligned}$$

  and the convex sets, for all $j \geq 1$,

$$\mathcal{A}^j = \{\alpha \in \mathcal{A} \;/\; \alpha H^{[k]} = 0, \;\; \text{for } 1 \leq k \leq j\;\}.$$

Summarily, Theorem 2.2 says that $agg(\alpha, P, \mathcal{B})$ is a homogeneous Markov chain if and only if for any possible past of the chain, with a last transition between two classes $B(l)$, $B(m)$ of $\mathcal{B}$, the conditional probability of type (1) associated with depends only on l and m. The set $\mathcal{A}^j$ has to be interpreted as regrouping all the initial distributions α for which, a possible past of size less than $j \geq 1$ (i.e. at most j classes are invoked in the conditional part of expression (1)) gives conditional probabilities of type (1) which depend only on the identity of the classes $B(l)$, $B(m)$. Consequently, the following representation of the set $\mathcal{A}_{\mathcal{M}}$ appears to be a natural one (see [10] for details.)

Theorem 2.6

$$\mathcal{A}_{\mathcal{M}} = \{\alpha \in \mathcal{A} \;/\; \alpha H^{[j]} = 0, \;\; j \geq 1\;\} \;=\; \bigcap_{j\geq 1} \mathcal{A}^j.$$

Consequently, if we construct the convex envelope of the family of vectors $\{\pi^{B(l)}$, $l \in F\}$, $\mathcal{A}_\pi = \sum_{l\in F} \lambda_l\, \pi^{B(l)}$ (with $\lambda_l \geq 0$ and $\sum_{l\in F} \lambda_l = 1$) and the convex set $\mathcal{A}_{\mathcal{M}}$ is non empty, then we have

$$\mathcal{A}_\pi \subseteq \mathcal{A}_{\mathcal{M}}.$$

It can be noted, as in [15], that the property of P-stability of $\mathcal{A}^j$ (i.e. $\mathcal{A}^j P \subseteq \mathcal{A}^j$) allows us to identify $\mathcal{A}_{\mathcal{M}}$ as the set $\mathcal{A}^j$. The example of Subsection 2.4 shows that the infinite intersection of $\mathcal{A}^j$'s can be finite and explicitly computed.

Finally, let us address the usual lumpability situation considered in the literature when any $\alpha \in \mathcal{A}$ leads to an aggregated homogeneous Markov chain $Y = agg(\alpha, P, \mathcal{B})$, i.e. when X is *strongly lumpable* with respect to the partition $\mathcal{B}$. We can characterize such a property of X by:

Theorem 2.7 *X is strongly lumpable with respect to the partition $\mathcal{B}$ iff for any couple of classes $B(l), B(m) \in \mathcal{B}$, $P(i, B(m))$ has the same value for all $i \in B(l)$. This common value represents the transition probability of going from state l to state m for the aggregated chain.*

2.2 Weak lumpability of R-positive Markov chains

We are now concerned with denumerable Markov chains with absorbing states which are assumed to be collapsed in only one class (state labeled by 0 for the aggregated process Y) of the partition $\mathcal{B}$. The other classes constitute a partition of the set of transient states, denoted by T, of X. It is easy to convince ourself that weak lumpability for such a Markov chain reduces to weak lumpability for the Markov chain with only one absorbing state and absorption probabilities equal to $P(i, B(0))$ for $i \in E$. Consequently, we consider only one absorbing state denoted by a (and $B(0) = \{a\}$.) Let us denote by $\mathcal{A}^T$ the subset of $\mathcal{A}$ composed of the distributions α with support T, i.e. $\sum_{i\in T} \alpha(i) = 1$. If we define the following subset of $\mathcal{A}_{\mathcal{M}}$

$$\mathcal{A}^T_{\mathcal{M}} \triangleq \{\alpha \in \mathcal{A}^T \ / \ agg(\alpha, P, \mathcal{B}) \text{ is a homogeneous Markov chain}\}$$

then we have

$$\mathcal{A}_{\mathcal{M}} = (1 - \lambda_T)\, 1^{\{a\}} + \lambda_T \mathcal{A}^T_{\mathcal{M}} \quad \text{where } 1 \geq \lambda_T \geq 0.$$

Therefore, we restrict the analysis to the set $\mathcal{A}^T_{\mathcal{M}}$.

In discrete time, the transition probability matrix P can be decomposed as follows:

$$P = \left(\begin{array}{c|c} 1 & 0 \\ \hline (I-Q)1^* & Q \end{array} \right),$$

where matrix Q is assumed to be irreducible. In this subsection, we recall (e.g. see Seneta [17, Chapter VI]) the definitions and the main properties of the R-classification of a non-negative irreducible matrix. It can be shown that all the power series $Q_{ij}(z) = \sum_{k=0}^{\infty} Q^k(i,j)z^k$, $i, j \in T$ have a common convergence radius, denoted by R, which is usually called the *convergence parameter* of matrix Q. If T is a finite set, then R is the inverse of the spectral radius of Q.

To address the asymptotic behaviour of matrices Q^k, we introduce in a similar way as in the usual stochastic case, the following "taboo" probabilities: $l^{(0)}_{ij} = f^{(0)}_{ij} = 0$ and for any $k \geq 1$, $i, j \in T$,

$$\left\{ \begin{array}{l} l^{(1)}_{ij} = Q(i,j) \\ l^{(k+1)}_{ij} = \sum_{r \neq i} l^{(k)}_{ir} Q(r,j) \end{array} \right. \qquad \left\{ \begin{array}{l} f^{(1)}_{ij} = Q(i,j) \\ f^{(k+1)}_{ij} = \sum_{r \neq j} Q(i,r) f^{(k)}_{rj}. \end{array} \right. \tag{3}$$

We recognize $l^{(k)}_{ij}$ as the transition probability of going from state i to state j in k steps without revisiting state i in the meantime. In the same way, $f^{(k)}_{ij}$ is the transition probability of going from state i to state j in k steps without visiting state j. We can define the power series :

$$L_{ij}(z) = \sum_k l^{(k)}_{ij} z^k \quad \text{and} \quad F_{ij}(z) = \sum_k f^{(k)}_{ij} z^k.$$

Since $f^{(k)}_{ij} \leq Q^k(i,j)$ and $l^{(k)}_{ij} \leq Q^k(i,j)$, these series are convergent at least for $|z| < R$.

Definition 2.8 *Let Q be an irreducible sub-stochastic matrix and $i \in T$:*

1. *state i is said to be R-transient if $F_{ii}(R-) < 1$;*
2. *state i is said to be R-recurrent if $F_{ii}(R-) = 1$;*
3. *a R-récurrent state i is R-positive or R-null according as*

$$\mu_i(R) \triangleq R\ F'_{ii}(R-) = \sum_k k f_{ii}^{(k)} R^k < \infty \text{ or } = \infty.$$

All the states share property 1, 2 *or* 3 *as soon as one of them has the property.*

For an R-recurrent matrix Q, there exists an unique (up to a constant) positive *R-invariant measure*, (resp. positive *R-invariant vector*) denoted by v (resp. w), that is

$$R\ vQ = v \qquad \text{(resp. } R\ Qw^* = w^* \text{)}.$$

We can now define the **stochastic** matrix $\overline{P}$ whose entries are given by

$$\overline{P}(i,j) \triangleq R\ \frac{w(j)}{w(i)}\ Q(i,j), \qquad i,j \in T.$$

Denoting the diagonal matrix with generic diagonal entry $w(i)$ by W, the previous relation becomes

$$\overline{P} = R\ W^{-1}\ Q\ W. \tag{4}$$

Denoting the sequence of "taboo" probabilities (3) associated with stochastic matrix $\overline{P}$ by $(\overline{f}_{ij}^{(k)})_{i,j\in T}$, it is straightforward to see that for any $i,j \in T$

$$\overline{f}_{ij}^{(k)} = R^k\ \frac{w(j)}{w(i)} f_{ij}^{(k)}.$$

From this last relation, we deduce the equality $\overline{\mu}_i \triangleq \sum_k k\,\overline{f}_{ii}^{(k)} = R(w(j)/w(i))\mu_i(R)$. Therefore, it is easy to show that matrix $\overline{P}$ is positive-recurrent if and only if matrix Q is R-positive. The stationary probability vector of $\overline{P}$ is $\pi = (v(i)w(i))_{i\in T}$ which gives a second characterization of the positive recurrence of $\overline{P}$: $\sum_i v_i w_i < \infty$. It is important to note that the R-recurrence property does not allow in any way to infer the convergence of the series $\sum_i v_i$ or $\sum_i w_i$.

It was shown in [9] that using quasi-stationary distribution can be fruitful for weak lumpability of a finite absorbing Markov chain. We propose in this subsection to extent some of those ideas to a R-positive Markov chain. A *quasi-stationary distribution* is a probability measure which makes stationary the following conditional probabilities :

$$\mathbb{P}_\alpha(X_n = i \mid X_n \in T) = \frac{(\alpha_T Q^n)(i)}{\alpha_T Q^n 1^*} \quad i \in T,$$

that is the vector $(\mathbb{P}_\alpha(X_n = i \,|\, X_n \in T))_{i\in T}$ is independent of n. It is easy to see that such a property is closely related to the existence of a R-invariant measure associated with matrix Q. It can be stated that R-recurrence is nearly a "minimal" assumption on matrix Q (up to Harrys Veech conditions, e.g. see Pruitt [12]) to consider such a measure. The existence of a quasi-stationary distribution under milder conditions than R-recurrence is discussed in recent papers (see e.g. Van Doorn [20], Kijima and Seneta [6],[7].) The R-positivity property of matrix Q is also the nearly "minimal" condition to have an ergodic interpretation of such a quasi-stationary distribution with **any** probability vector as initial distribution of the Markov chain X (see [10].) Moreover the results must include the finite state space ones reported in [9]. The following theorem gives an ergodic interpretation to the R-invariant measure v when it defines a probability distribution. Note that we don't make any distinction between periodic and aperiodic cases. Throughout the remainder of this subsection, we will assume that any initial distribution $\alpha \in \mathcal{A}^T$ satisfies a constraint of the type:

$$\alpha_T \le C_\alpha v \tag{5}$$

where C_α si a positive scalar. It allows to consider $\alpha_T W$ as a summable series because $0 < \alpha_T W 1^* \le C_\alpha \; vW1^* = C_\alpha \; \pi 1^* = C_\alpha$. Therefore the vector $(\alpha_T W)/(\alpha_T W 1^*)$ defines a probability distribution.

Lemma 2.9 *Let g be a non-negative function on T assumed to be π-integrable. For any initial distribution $\alpha \in \mathcal{A}^T$ with $\alpha_T \le C_\alpha v$, we have*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \left(\alpha_T W \overline{P}^k \right)(g) = \alpha_T W 1^* \; \pi(g).$$

Proof. From Result 2.3, we have that for any $i \in T$,

$$\lim_{n\to+\infty} \alpha_T W 1^* \; \frac{1}{n} \sum_{k=1}^{n} \left(\frac{\alpha_T W}{\alpha_T W 1^*} \overline{P}^k \right)(i) g(i) = \alpha_T W 1^* \; \pi(i) g(i).$$

Moreover, condition (5) required on α gives the following inequality for any $i \in T$:

$$\left(\alpha_T W \frac{1}{n} \sum_{k=1}^{n} \overline{P}^k \right)(i)\; g(i) \quad \le \quad C_\alpha \left(\frac{1}{n} \sum_{k=1}^{n} \pi \overline{P}^k \right)(i)\; g(i) = C_\alpha \; \pi(i) g(i).$$

Since $\sum_{i\in T} \pi(i) g(i) = \pi(g) < \infty$, the dominated convergence theorem allows us to write

$$\lim_{n\to\infty} \sum_{i\in T} \left(\alpha_T W \frac{1}{n} \sum_{k=1}^{n} \overline{P}^k \right)(i)\; g(i) \;=\; \sum_{i\in T} \alpha_T W 1^* \; \pi(i) g(i) \;=\; \alpha_T W 1^* \; \pi(g).$$

□

Theorem 2.10 *Let Q be a R-positive matrix such that its R-invariant measure v satisfies $v1^* < \infty$. Assume that α is a probability distribution which verifies relation (5). If we define the vector*

$$p_{n,\alpha} = \frac{\displaystyle\sum_{k=1}^{n} R^k \alpha_T Q^k}{\displaystyle\sum_{k=1}^{n} R^k \alpha_T Q^k 1^*}, \tag{6}$$

then we have

$$\lim_{n\to\infty} p_{n,\alpha} = \frac{v}{v1^*}.$$

This result can also be derived from Seneta and Vere-Jones (1966) but we use here standard arguments on regular Markov chains which give insight into the considered assumptions.

Proof. From definition (4) of matrix $\overline{P}$, we have

$$p_{n,\alpha} = \frac{(\alpha_T W \displaystyle\sum_{k=1}^{n} \overline{P}^k) W^{-1}}{(\alpha_T W \displaystyle\sum_{k=1}^{n} \overline{P}^k) W^{-1} 1^*}.$$

Let $g_j = W^{-1} e_j{}^* \geq 0$ and $g = W^{-1} 1^* \geq 0$, then we have $\pi(g_j) = v(j)$ and $\pi(g) = v1^* < \infty$. The Lemma 2.9 allows us to write for all initial distribution such that $\alpha_T \leq C_\alpha v$:

$$\lim_{n\to\infty} p_{n,\alpha} = \lim_{n\to\infty} \frac{(\alpha_T W \sum_{k=1}^{n} \overline{P}^k)(g_j)}{(\alpha_T W \sum_{k=1}^{n} \overline{P}^k)(g)} = \frac{\alpha_T W 1^* \; \pi(g_j)}{\alpha_T W 1^* \; \pi(g)} = \frac{v(j)}{v1^*}.$$

□

When the state space E is finite, it is clear that relation (5) is always satisfied and the convergence in Theorem 2.10 holds for any initial distribution. Under the assumptions of Theorem 2.10, we can derive an analogous result to Theorem 2.5. The proof is obtained with similar arguments, that is, firstly establishing the lemma

Lemma 2.11 *For any distribution $\alpha \in \mathcal{A}^T$ satisfying constraint (5), let β_n be the vector*

$$\beta_n = \frac{\displaystyle\sum_{k=1}^{n} R^k \alpha_T Q^k}{\displaystyle\sum_{k=1}^{n} R^k \alpha_T Q^k 1^*}.$$

Then, for all $l \neq 0$ and $m \in F$, we have

$$\begin{aligned}
\lim_{n\to\infty} f((0,\beta_n), B(l)) &= f((0,v), B(l)), \\
\lim_{n\to\infty} \mathbb{P}_{f((0,\beta_n),B(l))}(X_1 \in B(m)) &= \mathbb{P}_{f((0,v),B(l))}(X_1 \in B(m)).
\end{aligned}$$

Proof. The first limit is obtained by combining transformation (4) and Lemma 2.9 with, for $i \in T$, $g_j(i) = 1_{\{j\}}(i)/w(j)$ ($j \in B(l)$ is fixed) and $g(i) = 1_{B(l)}(i)/w(i)$. Since $f((0,\beta_n), B(l))(j) = \sum_{k=1}^n R^k \alpha_T Q^k(g)/\sum_{k=1}^n R^k \alpha_T Q^k(g)$, the limit, as n goes to infinity, is the ratio $v(j)/\sum_{i \in B(l)} v(i) = f((0,v), B(l))(j)$.

With the help of the previous limits, the second convergence derives also from transformation (4) and Lemma 2.9 with function g defined by:

$$g(i) = 1_{B(l)}(i) P(i, B(m))/w(i) \quad \forall i \in T.$$

Therefore, the two factors in the right-hand side of relation (2) (with the new vector β_n) tend respectively to $1/v_{B(l)}1^*$ and to $\sum_{i \in B(l)} v(i) P(i, B(m))$. $\square$

Let us define the set

$$\mathcal{A}^T_{\mathcal{M}}(v) \triangleq \{\alpha \in \mathcal{A}^T \;/\; \alpha_T \le C_\alpha v \text{ and } agg(\alpha, P, \mathcal{B}) \text{ is a homogeneous Markov chain}\}.$$

We are in position to show the following result.

Theorem 2.12 *Let X be a R-positive Markov chain with a R-invariant probability measure v (i.e. a quasi-stationary distribution.) If $\mathcal{A}^T_{\mathcal{M}}(v) \neq \emptyset$ then $(0, v) \in \mathcal{A}^T_{\mathcal{M}}(v)$. Moreover, if $\widehat{P}$ denotes the transition probability matrix of the homogeneous Markov chain $agg(\alpha, P, \mathcal{B})$ then this matrix is the same for all $\alpha \in \mathcal{A}^T_{\mathcal{M}}(v)$. The entries of matrix $\widehat{P}$ are given by*

$$\widehat{P}(l, m) = \sum_{i \in B(l)} v^{B(l)}(i)\; P(i, B(m)) \quad l \neq 0,\; m \in F \tag{7}$$

Proof. Let $\alpha \in \mathcal{A}^T$ satisfying (5) such that $agg(\alpha, P, \mathcal{B})$ is a homogeneous Markov chain with transition probability from state l to m denoted by $\widehat{P}(l, m)$. Let α_k be the vector

$$\alpha_k = \left(0, \frac{\alpha_T Q^k}{\alpha_T Q^k 1^*}\right).$$

For any k such that $\mathbb{P}_\alpha(X_k \in B(l)) > 0$ ($l \neq 0$), we have:

$$\begin{aligned}
\widehat{P}(l, m) &= \mathbb{P}_\alpha(X_{k+1} \in B(m) \mid X_k \in B(l)) \\
&= \mathbb{P}_{(\alpha_T Q^k)^{B(l)}}(X_1 \in B(m)) \\
&= \mathbb{P}_{{\alpha_k}^{B(l)}}(X_1 \in B(m)).
\end{aligned} \tag{8}$$

Choose n_0 large enough such that $\forall n \ge n_0$, $\sum_{k=1}^{n} (R^k \alpha_T Q^k)_{B(l)} 1^* > 0$ (T is irreducible). Let us denote by γ_k ($k = 1, \ldots, n$) the scalar

$$\gamma_k = \frac{(R^k \alpha_T Q^k)_{B(l)} 1^*}{\sum_{k=1}^{n} (R^k \alpha_T Q^k)_{B(l)} 1^*}.$$

The transition probability $\widehat{P}(l,m)$ can be rewritten

$$\begin{aligned}\widehat{P}(l,m) &= \sum_{\substack{1\le k\le n,\\ \mathbb{P}_\alpha(X_k\in B(l))>0}} \gamma_k\, \mathbb{P}_{{\alpha_k}^{B(l)}}(X_1\in B(m)) \qquad \text{with (8)}\\ &= \mathbb{P}_\Gamma(X_1\in B(m))\\ \text{where } \Gamma &= \sum_{\substack{1\le k\le n,\\ \mathbb{P}_\alpha(X_k\in B(l))>0}} \gamma_k(\alpha_k)^{B(l)} = {\beta_n}^{B(l)}.\end{aligned}$$

Therefore, we obtain

$$\widehat{P}(l,m) = \mathbb{P}_{f(\beta_n,B(l))}(X_1\in B(m)).$$

As n goes to infinity, we derive from Lemma 2.11 that the transition probabilities of the aggregated chain are (independent of α and) given by formula (7). □

The convexity of the set $\mathcal{A}^T_{\mathcal{M}}(v)$ follows from the unicity of the transition probability matrix for the aggregated chain.

Corollary 2.13 *If $\mathcal{A}^T_{\mathcal{M}}(v)\neq\emptyset$ then $\mathcal{A}^T_{\mathcal{M}}(v)$ is a convex set and it necessarily includes the convex subset $\mathcal{A}_v=\sum_{l\in F\setminus\{0\}}\lambda_l\,(0,v)^{B(l)}$ with $\lambda_l\ge 0$ and $\sum_{l\in F\setminus\{0\}}\lambda_l=1$.*

By definition, the set $\mathcal{A}^T_{\mathcal{M}}(v)$ is a subset of $\mathcal{A}^T_{\mathcal{M}}$. If $\mathcal{A}^T_{\mathcal{M}}(v)\neq\emptyset$ we trivially have $\mathcal{A}^T_{\mathcal{M}}\neq\emptyset$. The converse is true at least in the following specific cases which respectively ensure that $(0,v)\in\mathcal{A}^T_{\mathcal{M}}$.

Corollary 2.14 *If the set $\mathcal{A}^T_{\mathcal{M}}$ includes a distribution with finite support or if there is a class $B(l)$ ($l\neq 0$) within the partition $\mathcal{B}$, which collapses a finite number of states, then $(0,v)\in\mathcal{A}^T_{\mathcal{M}}$ and we have:*

$$\mathcal{A}^T_{\mathcal{M}}\neq\emptyset \Longleftrightarrow \mathcal{A}^T_{\mathcal{M}}(v)\neq\emptyset.$$

Proof. If $\mathcal{A}^T_{\mathcal{M}}$ includes a distribution α with finite support then we deduce from Lemme 2.11 and from the proof of Theorem 2.12 that $(0,v)\in\mathcal{A}^T_{\mathcal{M}}$.

Suppose that there exists one class $B(l)$, $l\neq 0$, with a finite number of states and that $\mathcal{A}^T_{\mathcal{M}}\neq\emptyset$. Let $\alpha\in\mathcal{A}^T_{\mathcal{M}}$, the irreducibility of transient states class T allows us to say that there exists $n\ge 1$ such that $(\alpha P^n)_{B(l)}1^*\neq 0$. Consequently, the distribution $(\alpha P^n)^{B(l)}$ belongs to $\mathcal{A}^T_{\mathcal{M}}$. The support of this distribution being finite, we deduce from the previous discussion that $(0,v)\in\mathcal{A}^T_{\mathcal{M}}(v)$.

□

Despite of restriction (5) on the initial distributions concerned with the previous theorem, the strong lumpability of R-positive Markov chains is characterized by the same statement (Theorem 2.7) as in the positive-recurrent case. The proposition is sufficient by the same arguments as in the proof of Theorem 2.7. It is necessary from the Corollary 2.14 applied to the initial distribution $1^{\{i\}}$ (finitly supported) for any $i\in E$.

2.3 Quasi-stationary distribution as a distribution of reset after absorption

In this subsection, we will show that the set $\mathcal{A}^T_{\mathcal{M}}(v)$ is non empty if and only if the set $\mathcal{A}_{\mathcal{M}}$ associated with a positive-recurrent chain is not empty too. Under the condition $v1^* = 1$, where v is the R-invariant measure of the R-positive matrix Q, we can define the following transition probability matrix denoted by $P^{(v)}$:

$$P^{(v)} = \begin{pmatrix} 0 & v \\ (I-Q)1^* & Q \end{pmatrix}.$$

Throughout this subsection, we carry on to denote the state associated with the first row of matrix $P^{(v)}$ by a in accordance with previous convention.

Lemma 2.15 *The Markov chain with transition probability matrix $P^{(v)}$ is irreducible and positive-recurrent. Its invariant probability measure is given by*

$$\pi^{(v)} = \lambda_R\, 1^{\{a\}} + (1-\lambda_R)(0,v) \quad \textit{with} \;\; \lambda_R = (R-1)/(2R-1). \tag{9}$$

Proof. The convergence parameter of the R-positive matrix Q is such that $R > 1$ (see [17].) Let us consider the “taboo” probability denoted by $f^{(k)}_{aa}$ and defined by the probability of going from state a to state a in k steps without revisiting state a in the meantime. The irreducible matrix $P^{(v)}$ will be recurrent if and only if $\sum_{k\geq 1} f^{(k)}_{aa} = 1$. Since v is R-invariant, we have $\sum_{k\geq 1} f^{(k)}_{aa} = \sum_{k\geq 1} vQ^{k-1}(I-Q)1^* = \sum_{k\geq 1}(1/R)^{k-1}(1-(1/R)) = 1$. Finally, the positive recurrence follows from checking that the invariant probability measure of matrix $P^{(v)}$ is given by formula (9). $\square$

We can now show the main result of this subsection.

Theorem 2.16 *If $\mathcal{A}_{\mathcal{M}}(P^{(v)})$ is the set of all initial distributions α such that $agg(\alpha, P^{(v)}, \mathcal{B})$ is a homogeneous Markov chain, we have $\mathcal{A}^T_{\mathcal{M}}(v) \neq \emptyset \iff \mathcal{A}_{\mathcal{M}}(P^{(v)}) \neq \emptyset$; in that case, we have*

$$\mathcal{A}^T_{\mathcal{M}}(v) \subseteq \mathcal{A}_{\mathcal{M}}(P^{(v)}) \subseteq \mathcal{A}_{\mathcal{M}}.$$

The transition probability matrix $\widehat{P^{(v)}}$ of $agg(\alpha, P^{(v)}, \mathcal{B})$ is given for every $m \in F$ by $\widehat{P^{(v)}}(l,m) = \widehat{P}(l,m)$ with $l \neq 0$ and by $\widehat{P^{(v)}}(0,m) = v_{B(m)}1^*$ (matrix $\widehat{P}$ is given by relation (7).)

Proof. The above one to one correspondence between the respective entries of matrices $\widehat{P}$ and $\widehat{P^{(v)}}$ is deduced from relation (9) in the previous lemma and from the definition of matrix $P^{(v)}$.

The inclusion of $\mathcal{A}_{\mathcal{M}}(P^{(v)})$ in $\mathcal{A}_{\mathcal{M}}$ follows in the same manner as in the finite case (see [9]) and is not reproduced here. We have only to prove that if $\mathcal{A}_{\mathcal{M}}(P^{(v)}) \neq \emptyset$ then $\mathcal{A}^T_{\mathcal{M}}(v) \neq \emptyset$. Indeed if $\mathcal{A}_{\mathcal{M}}(P^{(v)}) \neq \emptyset$ then $\pi^{(v)} \in \mathcal{A}_{\mathcal{M}}(P^{(v)}) \subseteq \mathcal{A}_{\mathcal{M}}$ from Theorem 2.5. It easily follows that $(\pi^{(v)})^T = (0,v) \in \mathcal{A}^T_{\mathcal{M}}$ and therefore, that $\mathcal{A}^T_{\mathcal{M}}(v) \neq \emptyset$.

The proposition ($\alpha \in \mathcal{A}^T_{\mathcal{M}}(v) \Longrightarrow \alpha \in \mathcal{A}^T_{\mathcal{M}}(P^{(v)})$) results directly from the proof of the inclusion $\mathcal{A}^T_{\mathcal{M}} \subseteq \mathcal{A}^T_{\mathcal{M}}(P^{(v)})$ in the finite case which can be found in [9].

□

We have already noted that $\mathcal{A}_{\mathcal{M}} = \lambda\ 1^{\{a\}} + (1-\lambda)\mathcal{A}^T_{\mathcal{M}}$ and that, in the finite case, $\mathcal{A}^T_{\mathcal{M}}(v) = \mathcal{A}^T_{\mathcal{M}}$. Therefore, the two sets $\mathcal{A}_{\mathcal{M}}$ and $\mathcal{A}_{\mathcal{M}}(P^{(v)})$ are identical. We are not able to establish the same equality in the denumerable case. Another important fact is that the two sets $\lambda\ 1^{\{a\}} + (1-\lambda)\mathcal{A}^T_{\mathcal{M}}(v)$ and $\mathcal{A}_{\mathcal{M}}(P^{(v)})$ are distinct in general (this will be illustrated in the example.) The equality will hold only in the case where any distribution in $\mathcal{A}_{\mathcal{M}}(P^{(v)})$ can be majorized by a multiple of the stationary distribution $\pi^{(v)}$ of $P^{(v)}$.

2.4 Example

Let us consider the following partition $\mathcal{B} = \{B(0) = \{0\}, B(1) = \{i \geq 1\}\}$ of the state space $E = \mathbb{N}$. The transition probability matrix P is given by:

$$\left(\begin{array}{l|ll}
P(0,0)=1 & P(0,1)=0 & P(0,n)=0 \text{ for } n\geq 2 \\ \hline
P(1,0)=0 & P(1,n)=(1/6)(5/6)^{n-1} & \text{for } n\geq 1 \\
P(n,0)=7/8 & P(n,1)=1/8 & P(k,n)=0 \\
\text{for any } n\geq 2 & \text{for } n\geq 2 & \text{for } k\geq 2,\ n\geq 2.
\end{array}\right).$$

The submatrix Q of transition probabilities between transient states (here, $T = B(1) = \{i \geq 1\}$) is clearly irreducible. We deduce that $\sum_{k\geq 1} f^{(k)}_{11} z^k = (1/6)z + (5/48)z^2$ (with the same notation as in the proof of Lemma 2.15.) It follows from [17, Def. 6.2] that state 1 is R-positive with $R = 12/5$ and therefore that all the transient states are R-positive too. The $12/5$-invariant probability measure v is given by

$$v(1) = \frac{1}{3}, \quad v(2) = \frac{1}{9}, \quad v(n) = \frac{1}{9}\left(\frac{5}{6}\right)^{n-1} \quad \forall n \geq 2.$$

We can directly check that $(0, v) \in \mathcal{A}^T_{\mathcal{M}}$. Indeed, we have only one transient class $B(1)$. The aggregated chain is a homogeneous Markov chain if and only if the distribution of the sojourn times in this class $B(1)$ is geometric with parameter $\widehat{P}(1,1) = 5/12$; this is immediate because vector v is precisely an $12/5$-invariant measure associated with matrix Q.

We compute the transition probability matrix $\widehat{P}$ associated with the aggregated chain $agg(\alpha, P, \mathcal{B})$ from relation (7)

$$\widehat{P} = \begin{pmatrix} 1 & 0 \\ 7/12 & 5/12 \end{pmatrix}.$$

If we form the irreducible recurrent positive matrix $P^{(v)}$ as in Subsection 2.3, then $agg(\alpha, P^{(v)}, \mathcal{B})$ has transition probability matrix $\widehat{P^{(v)}}$ (with Theorem 2.16)

$$\widehat{P^{(v)}} = \begin{pmatrix} 0 & 1 \\ 7/12 & 5/12 \end{pmatrix}.$$

According to Subsection 2.1, let us form matrices $\widetilde{P_a^{(v)}}$ and $\widetilde{P_1^{(v)}}$

$$\widetilde{P_a^{(v)}} = \Big(P^{(v)}(a, B(0)) \;\; P^{(v)}(a, B(1)) \Big) = (1\ 0)$$

$$\text{and} \quad \widetilde{P_1^{(v)}} = \begin{pmatrix} P^{(v)}(1, B(0)) & P^{(v)}(1, B(1)) \\ P^{(v)}(2, B(0)) & P^{(v)}(2, B(1)) \\ \vdots & \vdots \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 7/8 & 1/8 \\ \vdots & \vdots \end{pmatrix}.$$

Furthermore, the block diagonal matrix $H^{[1]} = \text{diag}(H_l)$ is constituted by $H_0 = \widetilde{P_a^{(v)}} - 1^* \widehat{P_a^{(v)}} = 0$ and

$$H_1 = \widetilde{P_1^{(v)}} - 1^* \widehat{P_1^{(v)}} = \begin{pmatrix} -7/12 & 7/12 \\ 7/24 & -7/24 \\ \vdots & \vdots \end{pmatrix}.$$

Convex $\mathcal{A}^1$ is defined by

$$\mathcal{A}^1 = \{\alpha \in \mathcal{A} / \;\; \alpha H^{[1]} = 0\}.$$

The linear system reduces to

$$\begin{aligned} \mathcal{A}^1 &= \{\alpha \in \mathcal{A} / \;\; \alpha_{B(1)} H_1 = 0\} \\ &= \{\alpha \in \mathcal{A} / \;\; -2\alpha(1) + \sum_{i \geq 2} \alpha(i) = 0\} \\ &= \{\alpha \in \mathcal{A} / \;\; 3\alpha(1) = 1 - \alpha(0)\}. \end{aligned}$$

We check now the $P^{(v)}$-stability of set $\mathcal{A}^1$, i.e. $\mathcal{A}^1 P^{(v)} \subseteq \mathcal{A}^1$. If $\alpha \in \mathcal{A}^1$ then we can write

$$(\alpha P^{(v)})(k) = \begin{cases} (1/4)\alpha(0) + (7/8)(\sum_{i\geq 2} \alpha(i)) & \text{for } k = 0, \\ (1/4)\alpha(0) + (1/6)\alpha(1) + (1/8)(\sum_{i\geq 2} \alpha(i)) & \text{for } k = 1, \\ (1/2)^n \alpha(0) + (1/6)(5/6)^{n-1}\alpha(1) & \text{for } k \geq 2. \end{cases}$$

From relation $3\alpha(1) = 1 - \alpha(0)$, it follows that

$$(\alpha P^{(v)})(k) = \begin{cases} 1/4 - \alpha(1) & \text{for } k = 0, \\ 1/3 + (1/3)\alpha(1) & \text{for } k = 1. \end{cases}$$

Finally $1 - (\alpha P^{(v)})(0) = 3/4 + \alpha(1) = 3(1/4 + (1/3)\alpha(1)) = 3(\alpha P)(1)$ and consequently $\alpha P^{(v)} \in \mathcal{A}^1$. The convex set $\mathcal{A}^1$ is stable by $P^{(v)}$ and we have $\mathcal{A}_{\mathcal{M}}(P^{(v)}) = \mathcal{A}^1$ as noted in Subsection 2.1.

We point out that $\mathcal{A}^T_{\mathcal{M}}(v) \subset \mathcal{A}_{\mathcal{M}}(P^{(v)})$. Indeed, choose $\alpha \in \mathcal{A}^T$ such that

$$\alpha(0) = 0, \; \alpha(1) = 1/3, \;\; \alpha(n) = \frac{2}{33}\left(\frac{11}{12}\right)^{n-1} \forall n \geq 2.$$

Since $3\alpha(1) = 1 - \alpha(0)$, we have $\alpha \in \mathcal{A}_{\mathcal{M}}(P^{(v)})$ but the ratio $\alpha(n)/v(n) \propto (11/10)^{n-1}$ is unbounded as n goes to infinity. Therefore, vector α cannot satisfy relation (5), so $\alpha \notin \mathcal{A}^T_{\mathcal{M}}(v)$.

3 Weak lumpability in continuous time

The weak lumpability property has been recently addressed in [1] for denumerable irreducible positive-recurrent Markov chains evolving in continuous time. The main result [1, Th. 2.3] is the counterpart of Theorem 2.5 in continuous time. Here, we propose to briefly discuss weak lumpability for denumerable Markov chains with the **only** assumption of having an uniform transition semi-group denoted by $(P_t)_{t\geq 0}$ (e.g. see Freedman (1983).) Let us recall some notions associated with continuous time Markov chains. A stochastic semi-group over the state space E, denoted by $(P_t)_{t\geq 0}$, is a family of matrices on E satisfying:

$$\begin{aligned}
&P_t \text{ is a stochastic matrix on } E \text{ for any } t \geq 0;\\
&\forall t, s \geq 0 \ : \ \ P_{t+s} = P_t\ P_s;\\
&P_0 = I.
\end{aligned}$$

The semi-group is said to be *standard* if

$$\lim_{t\to 0} P_t(i,i) = 1 \quad \forall i \in E.$$

We will assume that the semi-group is *uniform*, that is the previous limit is uniform in i. Such a semi-group is characterized by the following theorem.

Result 3.1 ([3, Th. 5.4.29]) *Let $(P_t)_{t\geq 0}$ be an uniform stochastic semi-group on E then $A = (dP_t/dt)_{t=0}$ exists and verifies*

$$\left.\begin{array}{l}
A(i,i) \leq 0 \ \ \forall i \in E,\\
A(i,j) \geq 0 \ \ \forall i \neq j,\\
\sup\{\ i : |A(i,i)|\ \} < \infty,\\
\sum_{j\in E} A(i,j) = 0 \ \ \forall i \in E.
\end{array}\right\} \tag{10}$$

Conversely, consider a matrix on E satisfying (10), there exists an unique uniform stochastic semi-group $(P_t)_{t\geq 0}$ such that $(dP_t/dt)_{t=0} = A$; namely

$$P_t = e^{At} \quad \forall t \geq 0.$$

Matrix A is called the generator of the semi-group.

It clearly appears that the generator A determines the semi-group $(P_t)_{t\geq 0}$ and that it is uniformly bounded. In particular, any finite Markov chain has an uniform transition semi-group. Consider a Poisson process $N = (N_t)_{t\geq 0}$ with rate a, such that $a \geq \sup(-A(i,i), i \in E)$. Let $(U_n)_{n\geq 0}$ be a discrete time Markov chain, with transition probability matrix U given by

$$U = I + A/a,$$

which is assumed to be independent of N. The process $(U_{N_t})_{t\geq 0}$ with transition semi-group given by

$$\sum_{n=0}^{\infty} e^{-at}\frac{(at)^n}{n!}U^n, \tag{11}$$

is stochastically equivalent to $X = (X_t)_{t\geq 0}$. The discrete-time Markov chain $(U_n)_{n\geq 0}$ is usually called the "uniformized" chain associated with $(X_t)_{t\geq 0}$. In [9], the result showing how to reduce the weak lumpability property from continuous time to discrete time is proved in the finite state space context. The proof is direct, avoiding preliminary works as in [16] (irreducible case.) Since the statement is only based on the definition of the Markov property and in the previous stochastic equivalence, this scheme still holds in the denumerable state space case and is given here for completeness.

Theorem 3.2 *Let X be a Markov chain with an uniform transition semi-group and generator A. The chain $agg(\alpha, A, \mathcal{B})$ is a homogeneous Markov chain iff $agg(\alpha, U, \mathcal{B})$ is also homogeneous Markov chain. So we have*

$$\mathcal{C}_{\mathcal{M}} \stackrel{\triangle}{=} \{\alpha \in \mathcal{A} \ / \ agg(\alpha, A, \mathcal{B}) \textit{ is a homogeneous Markov chain } \} = \mathcal{A}_{\mathcal{M}}(U).$$

Proof. For all $k \in \mathbb{N}$, $B_0, \ldots, B_k \in \mathcal{B}$, $0 < t_1 < \cdots < t_k$ and $0 < n_1 < \cdots < n_k$, we define, to simplify the notation,

$$F_X(k) = \mathbb{P}_\alpha(X_{t_k} \in B_k, \ldots, X_{t_1} \in B_1, X_0 \in B_0),$$

$$F_U(k) = \mathbb{P}_\alpha(U_{n_k} \in B_k, \ldots, U_{n_1} \in B_1, U_0 \in B_0),$$

$$F_N(k) = \mathbb{P}(N_{t_k} = n_k, \ldots, N_{t_1} = n_1).$$

Since N is a Poisson process with rate a, we have $F_N(k) > 0 \ \forall k \in \mathbb{N}$ and

$$F_N(k) = F_N(k-1)\ \mathbb{P}(N_{t_k - t_{k-1}} = n_k - n_{k-1}). \tag{12}$$

Probability $F_X(k)$ can be expressed in terms of the equivalent stochastic process U_{N_t} as:

$$\begin{aligned} F_X(k) &= \mathbb{P}_\alpha(U_{N_{t_k}} \in B_k, \ldots, U_{N_{t_1}} \in B_1, U_{N_{t_0}} \in B_0) \quad \forall k \geq 1 \\ &= \sum_{n_1 \geq 0} \sum_{n_2 \geq n_1} \cdots \sum_{n_{k-1} \geq n_{k-2}} \sum_{n_k \geq n_{k-1}} F_U(k)\ F_N(k) \qquad (13) \\ &\qquad \text{(from the independence of } U \text{ and of } N\text{)}\ . \end{aligned}$$

Assume that $agg(\alpha, U, \mathcal{B})$ is Markov homogeneous. This implies

$$F_U(k) = F_U(k-1)\ \mathbb{P}_\alpha(U_{n_k - n_{k-1}} \in B_k \mid U_0 \in B_{k-1}). \tag{14}$$

We have to show that

$$F_X(k) \quad = \quad F_X(k-1)\ \mathbb{P}_\alpha(X_{t_k - t_{k-1}} \in B_k \mid X_0 \in B_{k-1}).$$

Replacing $F_U(k)$ and $F_N(k)$ in (13) by the respective relations (12) and (14), we obtain

$$\begin{aligned}
F_X(k) &= \sum_{n_1 \geq 0} \cdots \sum_{n_k \geq n_{k-1}} F_U(k-1)\, \mathbb{P}_\alpha(U_{n_k - n_{k-1}} \in B_k \mid U_0 \in B_{k-1}) \\
&\qquad\qquad \times F_N(k-1)\, \mathbb{P}(N_{t_k - t_{k-1}} = n_k - n_{k-1}) \\
&= \sum_{n_1 \geq 0} \cdots \sum_{n_{k-1} \geq n_{k-2}} \sum_{l=0}^{+\infty} F_U(k-1) F_N(k-1) \\
&\qquad\qquad \times \mathbb{P}_\alpha(U_l \in B_k \mid U_0 \in B_{k-1})\, \mathbb{P}(N_{t_k - t_{k-1}} = l) \qquad (15) \\
&= \left(\sum_{n_1 \geq 0} \cdots \sum_{n_{k-1} \geq n_{k-2}} F_U(k-1)\, F_N(k-1) \right) \\
&\qquad\qquad \times \sum_{l \geq 0} \mathbb{P}_\alpha(U_l \in B_k \mid U_0 \in B_{k-1})\, \mathbb{P}(N_{t_k - t_{k-1}} = l) \\
&= F_X(k-1) \sum_{l=0}^{+\infty} \mathbb{P}_\alpha(U_l \in B_k \mid U_0 \in B_{k-1})\, \mathbb{P}(N_{t_k - t_{k-1}} = l),
\end{aligned}$$

that is,

$$F_X(k) = F_X(k-1)\, \mathbb{P}_\alpha(X_{t_k - t_{k-1}} \in B_k \mid X_0 \in B_{k-1}). \qquad (16)$$

Conversely assume that $agg(\alpha, A, \mathcal{B})$ is a homogeneous Markov chain. Relation (16) holds, so relation (15) holds too (since only formula (12) is invoked between these two expressions.) We can write

$$\begin{aligned}
F_X(k) &= \sum_{n_1 \geq 0} \cdots \sum_{n_{k-1} \geq n_{k-2}} \sum_{l \geq 0} F_U(k-1)\, F_N(k-1) \\
&\qquad\qquad \times \mathbb{P}_\alpha(U_l \in B_k \mid U_0 \in B_{k-1})\, \mathbb{P}(N_{t_k - t_{k-1}} = l) \\
&= \sum_{n_1 \geq 0} \cdots \sum_{n_k \geq n_{k-1}} F_U(k-1)\, \mathbb{P}_\alpha(U_{n_k - n_{k-1}} \in B_k \mid U_0 \in B_{k-1}) \\
&\qquad\qquad \times F_N(k-1)\, \mathbb{P}(N_{t_k - t_{k-1}} = n_k - n_{k-1}) \\
&= \sum_{n_1 \geq 0} \cdots \sum_{n_k \geq n_{k-1}} F_U(k-1)\, \mathbb{P}_\alpha(U_{n_k - n_{k-1}} \in B_k \mid U_0 \in B_{k-1})\, F_N(k)
\end{aligned}$$

with (12).

Using (13), we obtain the following relation:

$$\sum_{n_1=0}^{+\infty} \cdots \sum_{n_k=n_{k-1}}^{+\infty} F_N(k)\ \{F_U(k) - F_U(k-1)\mathbb{P}_\alpha(U_{n_k - n_{k-1}} \in B_k \mid U_0 \in B_{k-1})\} = 0.$$

Therefore, we deduce that for all $n_k > \cdots > n_1 > 0$,

$$F_U(k) = F_U(k-1)\, \mathbb{P}_\alpha(U_{n_k - n_{k-1}} \in B_k \mid U_0 \in B_{k-1}),$$

and so $agg(\alpha, U, \mathcal{B})$ is a homogeneous Markov chain. $\square$

Corollary 3.3 *If $\alpha \in \mathcal{C}_{\mathcal{M}}$ then the Markov chain $agg(\alpha, A, \mathcal{B})$ has a generator, denoted by $\widehat{A}$, which is given by $\widehat{A} = a(\widehat{U} - I)$ where $\widehat{U}$ is the transition probability matrix of $agg(\alpha, U, \mathcal{B})$.*

This result allows us to derive the unicity of the generator $\widehat{A}$ for all aggregated Markov chains under the assumptions of Theorems 2.5 or 2.12 for the (discrete time) "uniformized" chain $(U_n)_{n\geq 0}$. Specifically, if $(U_n)_{n\geq 0}$ is R-positive then the continuous time Markov chain $(X_t)_{t\geq 0}$ is λ-positive (in the terminology proposed by Kingman [8]) with $\lambda = a(1 - 1/R)$ (see Buiculescu [2] or [10].) Finally, we obtain

Corollary 3.4 *Let X be a Markov chain with an uniform transition semi-group and generator A.*

1. *Assume that X is irreducible positive-recurrent with invariant probability measure π. If $agg(\alpha, A, \mathcal{B})$ is a homogeneous Markov chain then it admits the generator $\widehat{A}$ given by*
$$\widehat{A}(l, m) = \sum_{i\in B(l)} \pi^{B(l)}(i)\, A(i, B(m)), \qquad \forall l, m \in F.$$

2. *Let X be a Markov chain with an irreducible transient class T and all its absorbing states are collapsed in the class $B(0)$ of the partition $\mathcal{B}$. The chain X is assumed to be λ-positive with a λ-invariant probability measure v. For any initial distribution α such that $\alpha_T \leq C_\alpha v$, where C_α is a positive real, if $agg(\alpha, A, \mathcal{B})$ is a homogeneous Markov chain, then its generator $\widehat{A}$ is given by*
$$\widehat{A}(l, m) = \sum_{i\in B(l)} v^{B(l)}(i)\, A(i, B(m)), \qquad \forall l \in F \setminus \{0\},\ \forall m \in F.$$

 Finally we have

$$\begin{aligned}\mathcal{C}_{\mathcal{M}}(v) &\stackrel{\triangle}{=} \{\alpha \in \mathcal{A} \ / \ \alpha_T \leq C_\alpha v \text{ and } agg(\alpha, A, \mathcal{B}) \text{ is a homogeneous Markov chain}\}\\ &= \{\alpha \in \mathcal{A} \ / \ \alpha_T \leq C_\alpha v \text{ and } agg(\alpha, U, \mathcal{B}) \text{ is a homogeneous Markov chain}\}.\end{aligned}$$

Under the conditions required by the previous corollary, characterization of strong lumpability can be deduced from their counterparts for discrete time Markov chains.

Corollary 3.5 *If the continuous time chain X is positive-recurrent or λ-positive with a λ-invariant probability measure, then X is strongly lumpable iff for any couple of classes $B(l)$, $B(m)$ in the partition $\mathcal{B}$, $A(i, B(m))$ has the same value for all $i \in B(l)$. This common value is the transition rate of going from state l to state m for the aggregated chain $agg(\alpha, A, \mathcal{B})$.*

We note that Corollary 2.14 may be expressed in the continuous time context. Another remark is that the first part of Corollary 3.4 is stated under milder conditions in [1, Th. 2.5]. We end the discussion pointing out that the equivalence between discrete time and continuous time using the uniformization technique, is also reported in Sumita and Rieders [19] for finite ergodic Markov chains. But it is based on an erroneous characterization given in [19, page 66,eq. (2.5)] of the weak lumpability property for a finite ergodic discrete time Markov chain:

the lumped chain $agg(\alpha, P, \mathcal{B})$ is a homogeneous Markov chain iff there exists a stochastic matrix $\widehat{P} = (\widehat{P}(l,m))_{l,m\in F}$ such that $\forall l, m \in F$:

$$\frac{\alpha_{B(l)}}{\alpha_{B(l)}1^*}(P^n)_{B(l)B(m)}1^* = \widehat{P}^n(l,m) \quad \forall n \geq 1. \tag{17}$$

Noting that the left hand side represents the probability $\mathbb{P}_\alpha(X_n \in B(m) \mid X_0 \in B(l))$, we see in fact that we require the Chapman-Kolmogorov condition on the transition probability matrix $\widehat{P}$ of the aggregated chain $agg(\alpha, P, \mathcal{B})$. This is generally false and famous counter-examples have been exhibited. One of the earliest has been given by Levy [11]. It is reformulated under various forms in Hachigan [4], Rosenblatt [13], ... In particular in [4] or [13] the non markovian chain whose transition probabilities satisfied Chapman-Kolmogorov condition is deduced from a state aggregation of a bi-dimensional Markov chain. Let us consider the finite aperiodic irreducible matrix P given in [13, Chap 3, Section1]:

$$P = \left(\begin{array}{cc|cc} 1/4 & 1/2 & 0 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ \hline 1/4 & 0 & 1/2 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{array}\right).$$

with stationary distribution $\pi = (1/4, 1/4, 1/4, 1/4)$. The partition is composed of $B(0) = \{1,2\}$ and $B(1) = \{3,4\}$ and the transition probability matrix associated with chain $agg(\alpha, P, \mathcal{B})$ given by Theorem 2.5 is

$$\widehat{P} = \begin{pmatrix} 5/8 & 3/8 \\ 3/8 & 5/8 \end{pmatrix}.$$

For all $n \geq 1$, we have

$$\widehat{P}^n = \begin{pmatrix} 1/2 + (1/2)(1/4)^n & 1/2 - (1/2)(1/4)^n \\ 1/2 - (1/2)(1/4)^n & 1/2 + (1/2)(1/4)^n \end{pmatrix}.$$

and we can verify that

$$P^n = \left(\begin{array}{cc|cc} 1/4 & 1/4 + (1/4)^n & 1/4 - (1/4)^n & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ \hline 1/4 & 1/4 - (1/4)^n & 1/4 + (1/4)^n & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{array}\right).$$

The following quantities are respectively $\widehat{P}^n(0,0)$, $\widehat{P}^n(0,1)$, $\widehat{P}^n(1,0)$, $\widehat{P}^n(1,1)$:

$$\begin{aligned}
\mathbb{P}_\pi(X_n \in B(0) \mid X_0 \in B(0)) = \frac{\pi_{B(0)}}{\pi_{B(0)}1^*}\,(P^n)_{B(0)B(0)}1^* &= \frac{1}{2}\,[1+(1/4)^n], \\
\mathbb{P}_\pi(X_n \in B(1) \mid X_0 \in B(0)) = \frac{\pi_{B(0)}}{\pi_{B(0)}1^*}\,(P^n)_{B(0)B(1)}1^* &= \frac{1}{2}\,[1-(1/4)^n], \\
\mathbb{P}_\pi(X_n \in B(0) \mid X_0 \in B(1)) = \frac{\pi_{B(1)}}{\pi_{B(1)}1^*}\,(P^n)_{B(1)B(0)}1^* &= \frac{1}{2}\,[1-(1/4)^n], \\
\mathbb{P}_\pi(X_n \in B(1) \mid X_0 \in B(1)) = \frac{\pi_{B(1)}}{\pi_{B(1)}1^*}\,(P^n)_{B(1)B(1)}1^* &= \frac{1}{2}\,[1+(1/4)^n].
\end{aligned}$$

The condition (17) is fulfilled for the initial distribution π.

We compute now

$$\begin{aligned}
\mathbb{P}_\pi(X_2 \in B(0),\ X_1 \in B(0),\ X_0 \in B(0)) &= \pi_{B(0)}(P_{B(0)B(0)})^2\,1^* \\
&= \frac{1}{4}(1,1)\begin{pmatrix} 1/4 & 1/2 \\ 1/4 & 1/4 \end{pmatrix}(3/4\;1/2)^* \\
&= 3/16.
\end{aligned}$$

If $agg(\pi, P, \mathcal{B})$ was a homogeneous Markov chain according to condition (17) then the probability $\mathbb{P}_\pi(X_2 \in B(0),\ X_1 \in B(0),\ X_0 \in B(0))$ will be equal to

$$(\pi_{B(0)}1^*)(\widehat{P}(0,0))^2 = \frac{25}{128}.$$

Consequently, the aggregated chain chain $agg(\pi, P, \mathcal{B})$ can not be markovian and we deduce from Theorem 2.5 that no initial distribution can lead to an aggregated markovian chain.

References

[1] Ball F. and Yeo G.F (1993), Lumpability and marginalisability for continuous-time Markov chains. *J. Appl. Prob.* **19**, 518–528.

[2] Buiculescu (1972), Quasi-stationary distributions on continuous Markov chains. *Revue Roumaine de Mathématiques Pures and Appliquées* **17**, 1013–1023.

[3] Freedman D. (1983), *Markov chains.* (Springer-Verlag).

[4] Hachigan J. (1963), Collapsed Markov chains and the Chapman-Kolmogorov equation. *Ann. Math. Statist.* **34**, 233–237.

[5] J.G. Kemeny and J.L Snell (1976), *Finite Markov chains*. (Springer-Verlag, New York).

[6] Kijima M. and Seneta E. (1991), Some results for quasi-stationary distributions of Birth-Death processes. *J. Appl. Prob.* **28**, 503–511.

[7] Kijima M. (1992), On the existence of quasi-stationary distributions in denumerable R-transient Markov Chains. *J. Appl. Prob.* **29**, 21–36.

[8] Kingman J.C. (1963), The exponential decay of Markov transition probabilities. *Proc. London Math. Soc.* **13**, 337–358.

[9] Ledoux J., Rubino G. and Sericola B. (1994), Exact aggregation of absorbing Markov processes using quasi-stationary distribution. *Forthcoming in J. Appl. Prob.*

[10] Ledoux J. (1993), *Markovian models: on the characterization of weak lumpability and on structural models for dependability of software.* (in French) Ph.D. thesis, University of Rennes I, 1993.

[11] Levy P. (1949), Exemples de processus pseudo-markoviens. *C. R. Acad. Sci. Paris* **228**(26), 2004–2006.

[12] Pruitt W.E. (1964), Eigenvalues of non-negative matrices. *Ann. Math. Statist.* **35**, 1797–1800.

[13] Rosenblatt M. (1971), *Markov Processes. Structure and Asymptotic Behavior*. (Springer-Verlag, New York).

[14] G. Rubino and B. Sericola (1989), On weak lumpability in Markov chains. *J. Appl. Prob.* **26**, 446–457.

[15] Rubino G. and Sericola B. (1991), A finite characterization of weak lumpable Markov processes. Part I: The discrete time case. *Stoch. Proc. and Appl.* **38**, 195-204.

[16] Rubino G. and Sericola B. (1993) A finite characterization of weak lumpable Markov processes Part II: The continuous time case. *Stoch. Proc. and Appl.* **45**, 115–125.

[17] Seneta E. (1981) *Non-negative matrices and Markov chains*. (Springer-Verlag, New York).

[18] Seneta E. and Vere-Jones D. (1966), On quasi-stationary distributions in discrete-time Markov chains with a denumerable infinity of states. *J. Appl. Prob.* **3**, 403–434.

[19] Sumita U. and Rieders M. (1989), Lumpability and time reversibility in the aggregation-disaggregation method for large Markov chains. *Comm. Statist.- Stochastic Models*, **5**, 63–81.

[20] E. Van Doorn (1991), Quasi-stationary distributions and Convergence to Quasi-stationarity of Birth-Death Processes. *Adv. Appl. Prob.*, **23**, 683–700.

Unité de recherche INRIA Lorraine, Technôpole de Nancy-Brabois, Campus scientifique,
615 rue de Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unité de recherche INRIA Rennes, IRISA, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 46 avenue Félix Viallet, 38031 GRENOBLE Cedex 1
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
ISSN 0249-6399